Variable-Fidelity Electromagnetic Simulations and Co-Kriging for Accurate Modeling of Antennas

Slawomir Koziel, *Senior Member, IEEE*, Stanislav Ogurtsov, Ivo Couckuyt, and Tom Dhaene, *Senior Member, IEEE*

***Abstract*—Accurate and fast models are indispensable in contemporary antenna design. In this paper, we describe the low-cost antenna modeling methodology involving variable-fidelity electromagnetic (EM) simulations and co-Kriging. Our approach exploits sparsely sampled accurate (high-fidelity) EM data as well as densely sampled coarse-discretization (low-fidelity) EM simulations that are accommodated into one model using the co-Kriging technique. By using coarse-discretization simulations, the computational cost of creating the antenna model is greatly reduced compared to conventional approaches, where high-fidelity simulations are directly used to set up the model. At the same time, the modeling accuracy is not compromised. The proposed technique is demonstrated using three examples of antenna structures. Comparisons with conventional modeling based on high-fidelity data approximation, as well as applications for antenna design, are also discussed.**

***Index Terms*—Antenna modeling, co-Kriging, computer-aided design, electromagnetic simulation, Kriging, surrogate modeling.**

I. Introduction

RELIABLE analytical or circuit models are only available for certain antenna structures. In general, full-wave electromagnetic (EM) analysis is the only way to evaluate the antenna reflection and radiation responses, particularly when interactions between an antenna and its environment (housing, feed lines, connectors, surrounding devices, etc.) have to be considered. Unfortunately, high-fidelity simulations are computationally expensive. This might be a serious obstacle for applying EM simulation tools for design tasks such as parametric design optimization or statistical analysis, where one has to evaluate the structure numerous times. With a long evaluation time, even traditional design approaches based on repetitive parameter sweeps turn into laborious procedures, not to mention attempts to automate the design process by linking the full-wave EM solver into the optimization loop where conventional algorithms (e.g., gradient-based ones) may require dozens or hundreds of objective function calls. Some design methods may require even a larger number of evaluations (e.g., statistical analysis or global optimization using evolutionary techniques [1]–[4]), which can be prohibitive from the point of view of the computational cost. Therefore, accurate and computationally cheap antenna models are indispensable.

Low-cost antenna models can be implemented using various approximation techniques, such as polynomial regression [5], radial basis functions [6], Kriging [7], [8], support vector regression [9]–[12], fuzzy systems [13], [14], artificial neural networks [15]–[18], or multidimensional Cauchy approximation [19]. Unfortunately, in order to ensure good accuracy over the entire design space, all of these techniques require a large number of training points. On the other hand, contemporary approximation models rely on uniform data sampling [7], which results in the exponential growth of the number of training points with respect to the dimensionality of the design space (the so-called *curse of dimensionality*). Consequently, approximation techniques are suitable for creating multiple-use library models rather than for *ad hoc* models for design tasks such as parametric optimization.

Another way of creating fast and reasonably accurate surrogates is to exploit physics-based low-fidelity models. The most popular technique of this kind in microwave engineering is space mapping (SM) [20]–[24]. In SM, the surrogate is constructed by means of a suitable correction of a low-fidelity (coarse) model of the microwave structure in question (high-fidelity or fine model), e.g., some auxiliary mappings applied to a circuit equivalent "reshape" the parameter space and/or response of the circuit [22]. The enhancement of the low-fidelity model is typically realized through suitable analytical formulas, which allows the surrogate model to be almost as computationally cheap as the coarse model. SM surrogate model identification is normally realized using a nonlinear parameter extraction process [22]. Due to the fact that the underlying coarse model embeds some knowledge about the structure under consideration, the accuracy of the SM surrogate is considerably better than the accuracy of possible function approximation models using a comparable amount of fine model data [22].

A drawback of SM models is that increasing the number of training points may have little effect on the model's quality [25], the main reason being the fact these models are effectively nonlinear regression models with a fixed number of parameters. Several ways of overcoming this problem have been proposed. SM modeling with variable weight coefficients [25] provides efficient utilization of available fine model data, however, at the expense of computational overhead related to a separate parameter extraction process required for each evaluation of the surrogate. SM modeling enhanced by fuzzy systems [26], radial

Manuscript received March 27, 2012; revised August 27, 2012; accepted November 16, 2012. Date of publication December 04, 2012; date of current version February 27, 2013. This work was supported in part by the Icelandic Centre for Research (RANNIS) Grant 120016021.. The work of I. Couckuyt is funded by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen).

S. Koziel and S. Ogurtsov are with the Engineering Optimization & Modeling Center, School of Science and Engineering, Reykjavik University, 101 Reykjavik, Iceland (e-mail: koziel@ru.is; stanislav@ru.is).

I. Couckuyt and T. Dhaene are with the Surrogate Modeling Laboratory, Department of Information Technology, Gent University—iMinds, B-9050 Gent, Belgium (e-mail: ivo.couckuyt@ugent.be; tom.dhaene@ugent.be).

Digital Object Identifier 10.1109/TAP.2012.2231924

basis functions [22], or Kriging [27] offer accuracy comparable with [25] without compromising computational cost; however, the implementation of these models is somewhat complicated.

Another issue of space mapping is that the underlying coarse model should be substantially faster than the fine model, because each evaluation of the SM surrogate requires evaluation of the corresponding coarse model. While some types of low-fidelity models are fast (e.g., analytical or circuit equivalent models), they have limited applicability for modeling antennas. The only versatile type of low-fidelity model available for all antennas can be obtained with coarse-discretization EM simulations. This type of low-fidelity model is, however, relatively expensive (typically, it is only 10 to 50 times faster than the high-fidelity simulation). This expense excludes SM from efficient modeling tools.

In this paper, we consider antenna models constructed using both high- and low-fidelity EM simulations. A simulation of a coarsely discretized antenna structure may not be very accurate; however, it is much faster than a high-fidelity simulation of the same antenna. As we demonstrate, such low-fidelity data can be combined with sparsely sampled high-fidelity simulation data using co-Kriging [28]. The resulting antenna model is as accurate as the conventional approximation surrogate, which uses a much larger number of training points. Also, once the model is established, neither the low- nor high-fidelity EM simulation needs to be launched any more.

Our modeling technique is demonstrated using three examples: an ultra-wideband (UWB) planar dipole antenna, a dielectric resonator antenna (DRA), and a wideband hybrid DRA. Comparison with the conventional Kriging interpolation is given. We also present the application of co-Kriging models for antenna optimization.

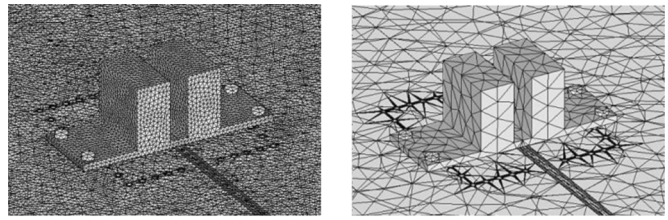
II. Antenna Modeling Using Co-Kriging

In this section, we discuss the construction of low- and high-fidelity EM models of antenna structures, describe the Kriging and co-Kriging interpolation techniques, as well as present the flow of the co-Kriging surrogate modeling of antennas.

A. Antenna Models

We consider two types of EM-simulation-based antenna models. Let $\boldsymbol{R}_f(\boldsymbol{x})$ denote an EM-simulated high-fidelity model, which is an accurate representation of the antenna structure [Fig. 1(a)]. $\boldsymbol{R}_f$ is expensive to evaluate, with the typical simulation time measured in hours. Here, $\boldsymbol{x}$ is a vector of designable (e.g., geometry or material) parameters. The components of $\boldsymbol{R}_f$ may represent, e.g., the antenna reflection coefficient $|S_{11}|$ over the frequency band of interest, a gain for a specified direction over the frequency band, etc. We also consider an auxiliary (low-fidelity) model $\boldsymbol{R}_c$, which may be evaluated using the same EM solver, however, with coarser discretization [Fig. 1(b)]. In practice, the low-fidelity model can be created not only by reducing the mesh density as compared to the high-fidelity one; other options may include the following:

- using smaller computational domain with the finite-volume discrete methods;
- using low-order basis functions with the finite-element and moment methods;



(a) (b)

Fig. 1. Dielectric resonator antenna [29] with different discretization density, an illustration: (a) a high-fidelity EM model with a fine tetrahedral mesh and (b) a low-fidelity EM model with a coarse tetrahedral mesh.

- using staircase shape approximation;
- applying simpler absorbing boundary conditions;
- exciting the computational domain with discrete sources rather than with full-wave ports;
- modeling metals with the perfect electric conductors;
- neglecting metallization thickness of traces, strips, and patches;
- ignoring dielectric losses and dispersion of materials;

Because of the possible simplifications listed above, the low-fidelity model $\boldsymbol{R}_c$ is (typically 10 to 50 times) faster than $\boldsymbol{R}_f$ but is not as accurate. Therefore, it cannot be normally directly used instead of the high-fidelity model to perform tasks such as design optimization. In this paper, we combine the low- and high-fidelity simulations to create the surrogate model that is almost as accurate as $\boldsymbol{R}_f$ but requires much smaller number of high-fidelity training points than the approximation model created using only $\boldsymbol{R}_f$ samples.

B. Kriging Interpolation

A well-known technique in surrogate modeling is Kriging [7], [43]. Kriging surrogate models are also known as Gaussian processes (GPs) [32] or Gaussian random fields [45]. Originally proposed by Krige [44], Kriging was popularized for the Design and Analysis of Computer Experiments (DACE) by Sacks *et al.* [30], where it has proven to be very useful for tasks such as optimization [33], [34], design space exploration, visualization, prototyping, and sensitivity analysis [31], [35]. For a full survey of Kriging, the reader is referred to [32] and [43]. In this section, a summary is given of the most important aspects of Kriging.

Let $X_{\mathrm{B.KR}} = (\boldsymbol{x}^1_{\mathrm{KR}}, \boldsymbol{x}^2_{\mathrm{KR}}, \ldots, \boldsymbol{x}^{N_{\mathrm{KR}}}_{\mathrm{KR}})$ be the base (training) set and $\boldsymbol{R}_f(X_{\mathrm{B.KR}})$ the associated fine model responses. Kriging first fits a regression function on the data and, subsequently, constructs a GP through the residuals. The idea is that the regression function captures the largest variance in the data, while the GP takes care of the finer details and the final interpolation. This is reflected in the Kriging interpolant, which is derived as

$$\boldsymbol{R}_{\mathrm{s.KR}}(\boldsymbol{x}) = M\alpha + r(\boldsymbol{x}) \cdot \Psi^{-1} \cdot (\boldsymbol{R}_f(X_{\mathrm{B.KR}}) - F\alpha) \quad (1)$$

where M and F are model matrices of the test point $\boldsymbol{x}$ and the base set $X_{\mathrm{B.KR}}$, respectively, representing the regression function. The coefficients of the regression function, i.e., the vector α, are determined by generalized least squares,

$$\alpha = (X'_{\mathrm{B.KR}}\Psi^{-1}X_{\mathrm{B.KR}})^{-1}X_{\mathrm{B.KR}}\Psi^{-1}\boldsymbol{R}_f(X_{\mathrm{B.KR}})$$

$r(\boldsymbol{x}) = (\Psi(\boldsymbol{x}, \boldsymbol{x}_{\mathrm{KR}}^1), \ldots, \Psi(\boldsymbol{x}, \boldsymbol{x}_{\mathrm{KR}}^{N_{\mathrm{KR}}}))$ is an $1 \times N_{\mathrm{KR}}$ vector of correlations between the point $\boldsymbol{x}$ and the base set $X_{\mathrm{B.KR}}$, and Ψ is a $N_{\mathrm{KR}} \times N_{\mathrm{KR}}$ correlation matrix given by

$$\Psi = \begin{bmatrix} \Psi\left(\boldsymbol{x}_{\mathrm{KR}}^1, \boldsymbol{x}_{\mathrm{KR}}^1\right) & \cdots & \Psi\left(\boldsymbol{x}_{\mathrm{KR}}^1, \boldsymbol{x}_{\mathrm{KR}}^{N_{\mathrm{KR}}}\right) \\ \vdots & \ddots & \vdots \\ \Psi\left(\boldsymbol{x}_{\mathrm{KR}}^{N_{\mathrm{KR}}}, \boldsymbol{x}_{\mathrm{KR}}^1\right) & \cdots & \Psi\left(\boldsymbol{x}_{\mathrm{KR}}^{N_{\mathrm{KR}}}, \boldsymbol{x}_{\mathrm{KR}}^{N_{\mathrm{KR}}}\right) \end{bmatrix}.$$

Moreover, Kriging also predicts the approximation error (prediction variance) at each location in the design space, also where no simulations has been chosen yet. The approximation error is zero in the data points themselves, as Kriging interpolates all data.

The regression function actually operates as the mean of the GP: predictions too far from existing data (e.g., outside the sampled region) will revert to the mean (=regression function). As the behavior of the response is usually unknown, a popular choice that works well is the constant regression function, also known as *ordinary* Kriging. However, in this case, Kriging is purely an interpolation technique (in contrast to extrapolation). By using prior knowledge or other techniques (e.g., blind Kriging [46]), one may identify basis functions (linear, quadratic, etc.) to use in the regression function, enabling Kriging to extrapolate outside the sampled region. This is especially useful for problems with missing data, i.e., large gaps in the input space where no data is available.

Arguably, the choice of correlation function is crucial to create an accurate Kriging surrogate model. A popular class of correlation functions is defined by $\psi(\boldsymbol{x}, \boldsymbol{x}') = \exp(\Sigma_{k=1,\ldots,n} - \theta_k |x^k - x'^k|^p)$. These correlation functions are called *stationary* since the correlation function only depends on the distance between the two points $\mathbf{x}$ and $\mathbf{x}'$. The smaller the distance between two points, the higher the correlation and, hence, the more the prediction of one point is influenced by the other. Similarly, if the distance increases the correlation drops to zero.

The rate and manner at which this happens is governed by several parameters. In essence, the parameter p determines the "smoothness" of the prediction; see Fig. 2(a). A value of $p = 2$ leads to a smooth prediction, but also has strict smoothness requirements on the response $\boldsymbol{R}_f$. With a smaller value of p, the correlation decreases much faster as the two points move farther from each other, which is suitable for more sharp (discontinuous) responses.

The second set of parameters, $\theta_1, \ldots, \theta_n$, describes the influence sphere of a point on nearby points for each dimension; see Fig. 2(b). This is useful as it describes the linearity of the response and, hence, can be used to identify relevant variables.

Usually, p is set fixed while the parameters $\theta_1, \ldots, \theta_n$ are identified using maximum-likelihood estimation (MLE) [37]. In particular, we minimize the negative concentrated log-likelihood,

$$\ln(L) \cong -\frac{N_{\mathrm{KR}}}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln(|\Psi|),$$

where

$$\hat{\sigma}^2 = (\boldsymbol{R}_f(X_{\mathrm{B.KR}}) - F\alpha)' \Psi^{-1} (\boldsymbol{R}_f(X_{\mathrm{B.KR}}) - F\alpha)/N_{\mathrm{KR}}.$$

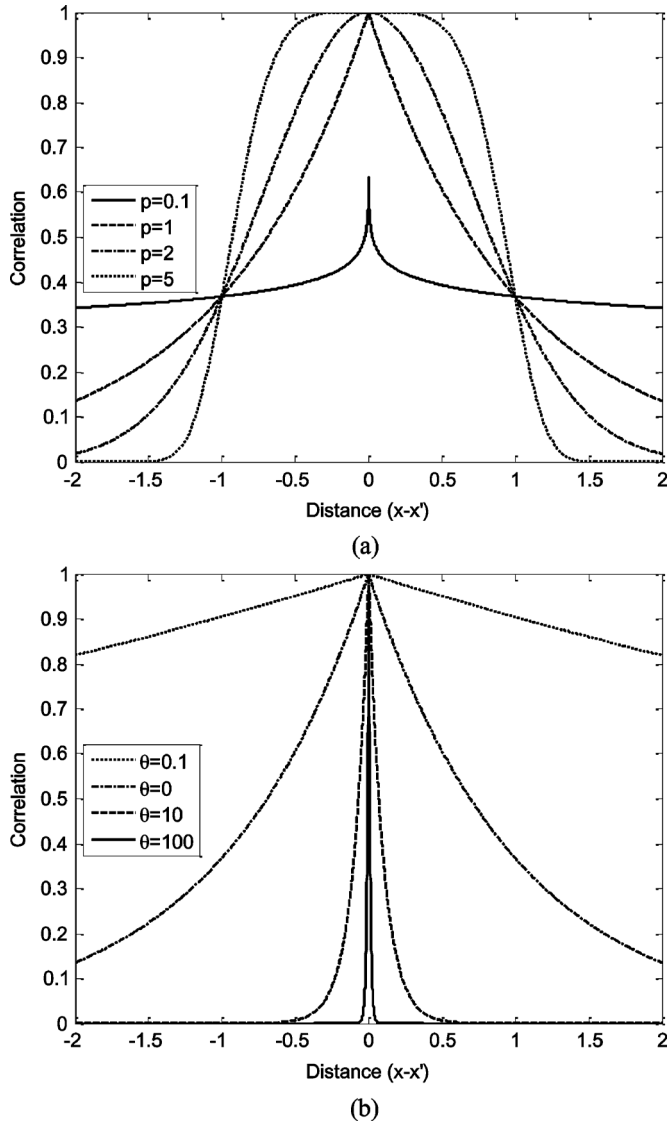


Fig. 2. Examples of one-dimensional correlation functions: (a) with varying parameter p for $\theta = 1$ (b) with varying parameter θ for $p = 1$.

Often, the parameter p is set to two, also known as the Gaussian correlation function, which is suitable for many problems. However, the problems discussed in this paper have a rather sharp response and, thus, it is chosen to use $p = 1$, also known as the exponential correlation function. Lastly, as we do not need any extrapolation capabilities, the regression function is set constant, i.e., $F = [1\ 1\ \ldots\ 1]^T$ and $M = (1)$.

Note that the application of Kriging is primarily limited by the size of the dataset. The number of samples has a direct impact on the correlation matrix Ψ, which grows quickly as the number of samples increases. As the inverse of Ψ needs to be computed many times during the MLE, Kriging is typically used for datasets with less than 1000 samples for which the computation time is minutes (just seconds for <100 samples), although it is possible to apply Kriging to datasets up to 5000 samples or more, depending on the available memory. Evaluation of a Kriging model does not have this problem and is usually in the order of milliseconds even for thousands of points.

C. Co-Kriging Interpolation

The popularity of Kriging has generated a large body of research, including several extensions to Kriging to handle different problem settings, e.g., by adding gradient information in the prediction [37], or by approximating stochastic simulations [36], etc. Co-Kriging is another interpolation flavor, which exploits the correlation between fine and coarse model data to enhance the prediction accuracy [38]. In this paper the autoregressive co-Kriging model of Kennedy *et al.* [28] is adopted.

Creating a co-Kriging model can be interpreted as constructing two Kriging models in sequence. First a Kriging model $\boldsymbol{R}_{\mathrm{s.KRc}}$ of the coarse data $(X_{\mathrm{B.KRc}}, \boldsymbol{R}_c(X_{\mathrm{B.KRc}}))$ is constructed. Subsequently, the second Kriging model $\boldsymbol{R}_{\mathrm{s.KRd}}$ is constructed on the residuals of the fine and coarse data $(X_{\mathrm{B.KRf}}, \boldsymbol{R}_d)$, where $\boldsymbol{R}_d = \boldsymbol{R}_f(X_{\mathrm{B.KRf}}) - \rho \cdot \boldsymbol{R}_c(X_{\mathrm{B.KRf}})$. The parameter ρ is included in the MLE of the second Kriging model. If the response values $\boldsymbol{R}_c(X_{\mathrm{B.KRf}})$ are not available, they can be approximated by the first Kriging model $\boldsymbol{R}_{\mathrm{s.KRc}}$, namely, $\boldsymbol{R}_c(X_{\mathrm{B.KRf}}) \approx \boldsymbol{R}_{\mathrm{s.KRc}}(X_{\mathrm{B.KRf}})$.

Note that the configuration (choice of correlation function, regression function, etc.) of both Kriging models can be adjusted separately for the coarse data R_f and the residuals R_d, respectively.

The final co-Kriging model $\boldsymbol{R}_{\mathrm{s.CO}}$ is built upon the two Kriging models. Namely, the co-Kriging interpolant is defined similarly as (1),

$$\boldsymbol{R}_{\mathrm{s.CO}}(\boldsymbol{x}) = M\alpha + r(\boldsymbol{x}) \cdot \Psi^{-1} \cdot (\boldsymbol{R}_d - F\alpha) \tag{2}$$

where the block matrices $M, F, r(\boldsymbol{x})$ and Ψ can be written as a function of the two underlying Kriging models $\boldsymbol{R}_{\mathrm{s.KRc}}$ and $\boldsymbol{R}_{\mathrm{s.KRd}}$:

$$\begin{aligned} r(\boldsymbol{x}) &= \left[\rho \cdot \sigma_c^2 \cdot r_c(\boldsymbol{x}), \rho^2 \cdot \sigma_c^2 \cdot r_c(\boldsymbol{x}, X_{\mathrm{B.KR}_f}) + \sigma_d^2 \cdot r_d(\boldsymbol{x})\right] \\ \Psi &= \begin{bmatrix} \sigma_c^2 \Psi_c & \rho \cdot \sigma_c^2 \cdot \Psi_c(X_{\mathrm{B.KR}_c}, X_{\mathrm{B.KR}_f}) \\ 0 & \rho^2 \cdot \sigma_c^2 \cdot \Psi_c(X_{\mathrm{B.KR}_f}, X_{\mathrm{B.KR}_f}) + \sigma_d^2 \cdot \Psi_d \end{bmatrix} \\ F &= \begin{bmatrix} F_c & 0 \\ \rho \cdot F_d & F_d \end{bmatrix}, \quad M = [\rho \cdot M_c \; M_d] \end{aligned} \tag{3}$$

where $(F_c, \sigma_c, \Psi_c, M_c)$ and $(F_d, \sigma_d, \Psi_d, M_d)$ are matrices obtained from the Kriging models $\boldsymbol{R}_{\mathrm{s.KRc}}$ and $\boldsymbol{R}_{\mathrm{s.KRd}}$, respectively (see Section II-B). In particular, σ_c^2 and σ_d^2 are process variances, while $\Psi_c(\cdot,\cdot)$ and $\Psi_d(\cdot,\cdot)$ denote correlation matrices of two datasets with the optimized $\theta_1, \ldots, \theta_n$ parameters and correlation function of the Kriging models $\boldsymbol{R}_{\mathrm{s.KRc}}$ and $\boldsymbol{R}_{\mathrm{s.KRd}}$, respectively. The block matrix Ψ is the crucial part of co-Kriging, as it is here that the correlation between the coarse and fine model data is taken into account.

Similarly to Section II-B, we choose the exponential correlation function and the constant regression function for the underlying Kriging models, $\boldsymbol{R}_{\mathrm{s.KRc}}$ and $\boldsymbol{R}_{\mathrm{s.KRd}}$.

For illustration purposes, Kriging and co-Kriging are applied to a mathematical example; see Fig. 3. Using the same fine model data, co-Kriging is able to capture the behavior of the fine model better than Kriging, which is attributed to the use of the additional coarse model data.

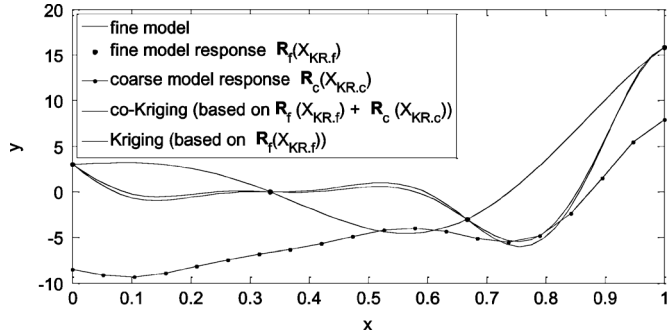


Fig. 3. Kriging and co-Kriging applied to a 1D mathematical example. Co-Kriging interpolates the fine model response and is further corrected by the coarse model response.

D. Co-Kriging Surrogate Modeling Flow

Co-Kriging models are created in a region of interest $X = [\boldsymbol{x}^0 - \boldsymbol{\delta}, \boldsymbol{x}^0 + \boldsymbol{\delta}]$, where $\boldsymbol{x}^0 \in R^n$ is a center and $\boldsymbol{\delta} \in R^n$ is a region size. The modeling procedure can be summarized as follows (where N_{KRc} and N_{KRf} are the numbers of the low- and high-fidelity model samples used to set up the model).

1. Allocate N_{KRc} samples $X_{\mathrm{B.KRc}} = \{\boldsymbol{x}^1_{\mathrm{KRc}}, \boldsymbol{x}^2_{\mathrm{KRc}}, \ldots, \boldsymbol{x}^{\mathrm{NKRc}}_{\mathrm{KRc}}\}$ using the modified Latin hypercube sampling (LHS) [40].
2. Evaluate the low-fidelity model to get the low-fidelity training set $(X_{\mathrm{B.KRc}}, \boldsymbol{R}_c(X_{\mathrm{B.KRc}}))$.
3. Allocate N_{KRf} samples $X_{\mathrm{B.KRf}} = \{\boldsymbol{x}^1_{\mathrm{KRf}}, \boldsymbol{x}^2_{\mathrm{KRf}}, \ldots, \boldsymbol{x}^{\mathrm{NKRf}}_{\mathrm{KRf}}\}$ using LHS [39].
4. Evaluate the high-fidelity model to get the high-fidelity training set $(X_{\mathrm{B.KRf}}, \boldsymbol{R}_f(X_{\mathrm{B.KRf}}))$.
5. Construct the co-Kriging surrogate model $\boldsymbol{R}_{\mathrm{s.CO}}$ as in (2) and (3). Finally, once constructed, the co-Kriging model does not refer to EM simulation.

Construction of the co-Kriging surrogate model, once the data point from both low- and high-fidelity models are available, is relatively fast and takes from a fraction of a second (if N_{KRc} and N_{KRf} equal 20 to 50) to a few seconds per frequency (if the number of training points is a few hundred). In a typical situation (e.g., 100 frequency points), the model construction typically takes several minutes. Evaluation of the co-Kriging model is very fast and takes a fraction of a second.

III. Verification Examples

In this section, we present numerical verification of the co-Kriging surrogate modeling technique of Section II. In order to reduce the computational cost of creating the surrogate and exploiting the knowledge about the antenna structure contained in the low-fidelity model, the number of low-fidelity model samples N_{KRc} should be substantially smaller than the number of high-fidelity model samples N_{KRf} (cf. Section II-D). However, for verification purposes, we consider various setups with the fixed $N_{\mathrm{KRc}} = 400$ and various N_{KRf} ranging from 20 to 400. The co-Kriging model is compared with Kriging interpolation of high-fidelity model data, also for different sizes of training sets from 20 to 400 samples.

A. UWB Planar Dipole Antenna

Consider the planar dipole antenna [40] (Fig. 4). The design variables are $\boldsymbol{x} = [l_0 \; w_0 \; a_0 \; l_p \; w_p \; s_0]^T$. The high-fidelity model

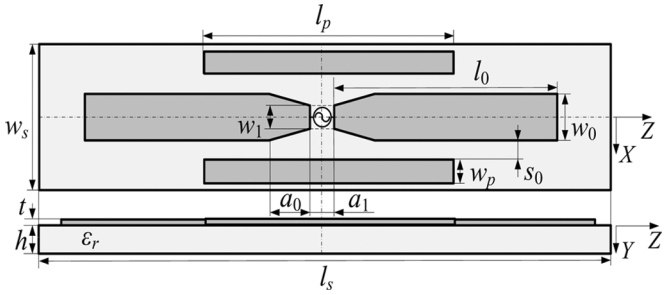


Fig. 4. Dipole antenna geometry [40]: top and side views. The dashed–dotted lines show the magnetic (YOZ) and the electric (XOY) symmetry walls. A 50-Ω source feeds the dipole.

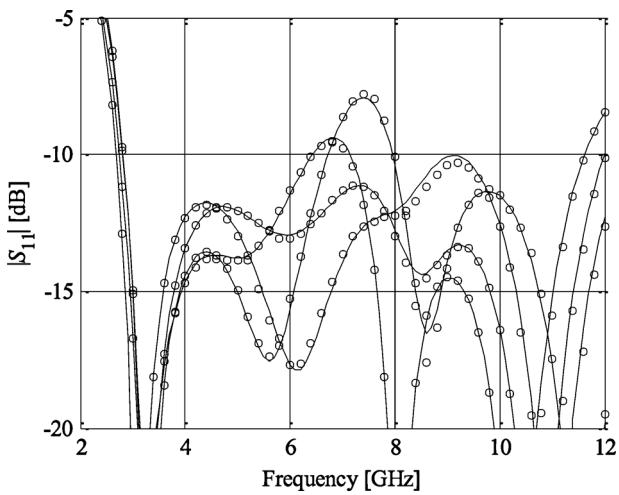


Fig. 5. UWB dipole: responses of $\boldsymbol{R}_f$ (—) and co-Kriging surrogate model (o) at selected test points. Co-Kriging model created using 50 evaluations of $\boldsymbol{R}_f$ and 400 evaluations of $\boldsymbol{R}_c$.

TABLE I
UWB DIPOLE ANTENNA: MODELING RESULTS

Model	Average Modeling Error [%]				
	$N_{KR}=20$	$N_{KR}=50$	$N_{KR}=100$	$N_{KR}=200$	$N_{KR}=400$
$\boldsymbol{R}_{s.KR}$	17.5	5.6	4.3	2.8	2.0
$\boldsymbol{R}_{s.CO}$	4.2	2.6	2.4	2.0	1.9

$\boldsymbol{R}_f$ (∼10 million mesh cells) is simulated using the CST MWS transient solver [41] in 44 min. The low-fidelity model $\boldsymbol{R}_c$ is also simulated with the same CST MWS solver (∼100 000 mesh cells, evaluation time of 43 s). Both antenna models are set up in the region with the center at $\boldsymbol{x}^0 = [19\ 13\ 0.5\ 13\ 6\ 1]^T$ and size $\boldsymbol{\delta} = [1\ 1\ 0.2\ 1\ 1\ 0.2]^T$. The Kriging and co-Kriging models $(\boldsymbol{R}_{\mathrm{s.KR}}, \boldsymbol{R}_{\mathrm{s.CO}})$ are constructed using various numbers of training points (from $N_{\mathrm{KR}} = 20$ to $N_{\mathrm{KR}} = 400$). Co-Kriging models are configured using 400 $\boldsymbol{R}_c$ samples (the CPU cost of which corresponds to around six evaluations of $\boldsymbol{R}_f$). The quality of the surrogate is assessed using a relative error measure $\|\boldsymbol{R}_f(\boldsymbol{x}) - \boldsymbol{R}_s(\boldsymbol{x})\|/\|\boldsymbol{R}_f(\boldsymbol{x})\|$ expressed in percent. The error is averaged over 50 test designs.

The modeling errors are given in Table I (see also Fig. 5). Note that the co-Kriging model accuracy obtained with 20 (50) $\boldsymbol{R}_f$ samples is as good as that of the Kriging model obtained for 100 (200) samples, which proves that the co-Kriging technique and the use of coarse-discretization EM data allows us to greatly reduce the computational cost of creating an accurate antenna model compared to conventional methods based solely on $\boldsymbol{R}_f$ data.

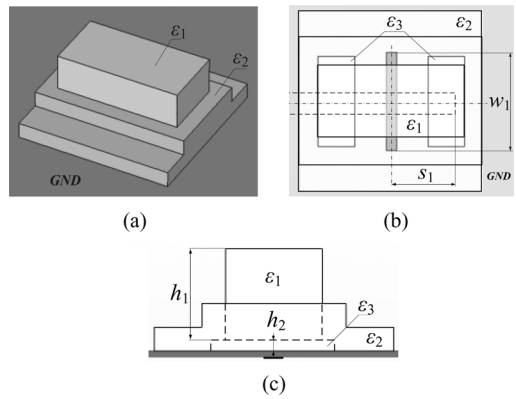


Fig. 6. DRA geometry [42]: (a) 3D view; (b) top view; and (c) front view.

As mentioned in Section II-D, creation of the co-Kriging model only takes several minutes; the major component of the computational cost is acquiring the training data. For this example, evaluation of the training points for the co-Kriging model with $\mathrm{N_{KRc}} = 400$ and $\mathrm{N_{KRf}} = 20$ took about 20 h, which is almost four times less than gathering data for the Kriging model with $\mathrm{N_{KR}} = 100$, which has the same accuracy.

B. Rectangular Dielectric Resonator Antenna

Consider a suspended rectangular DRA [42] (Fig. 6). The DRA is energized with a 50–Ω microstrip line through a slot made in the ground plane. The design variables are $\boldsymbol{x} = [\varepsilon_1\ h_1\ h_2\ s_1\ w_1]^T$. Other parameters are fixed. The high- and low-fidelity models are evaluated in CST MWS [41] with the following evaluation times: $\boldsymbol{R}_f$ in 11 min, and $\boldsymbol{R}_c$ in 20 s. Both antenna models are set up in the region with the center at $\boldsymbol{x}^0 = [10\ 8.5\ 0.5\ 3\ 10]^T$ and size $\boldsymbol{\delta} = [1\ 1\ 0.5\ 1\ 2]^T$ mm.

Qualitatively, the results are similar to those for the previous example, i.e., co-Kriging allows us to substantially reduce the computational cost of creating the accurate antenna model when compared to approximation of the high-fidelity model data only (cf. Table II and Fig. 7).

C. Wideband Hybrid Dielectric Resonator Antenna

As the last example, consider a wideband hybrid antenna [42], shown in Fig. 8. A quarter-wavelength monopole is loaded by dielectric ring resonator. The design variables are $\boldsymbol{x} = [h_1\ h_2\ r_1\ r_2\ g]^T$. Other parameters are fixed. The high- and low-fidelity models are evaluated in CST MWS [41] with the following evaluation times: $\boldsymbol{R}_f$ in 16 minutes, and $\boldsymbol{R}_c$ in 30 s. Both antenna models are set up in the region with the center at $\boldsymbol{x}^0 = [4\ 10\ 2\ 4\ 0.5]^T$ and size $\boldsymbol{\delta} = [0.5\ 1.0\ 0.5\ 0.5\ 0.5]^T$ mm.

The results are consistent with the previous examples. In particular, co-Kriging allows accuracy comparable to Kriging interpolation of the high-fidelity model data with substantially smaller number of training data samples (cf. Table III and Fig. 9).

IV. APPLICATION EXAMPLES: ANTENNA OPTIMIZATION

The co-Kriging surrogate models have been also applied to optimize the antenna structures considered in Section III. Here, the specific setup was $N_{\mathrm{KRc}} = 400$

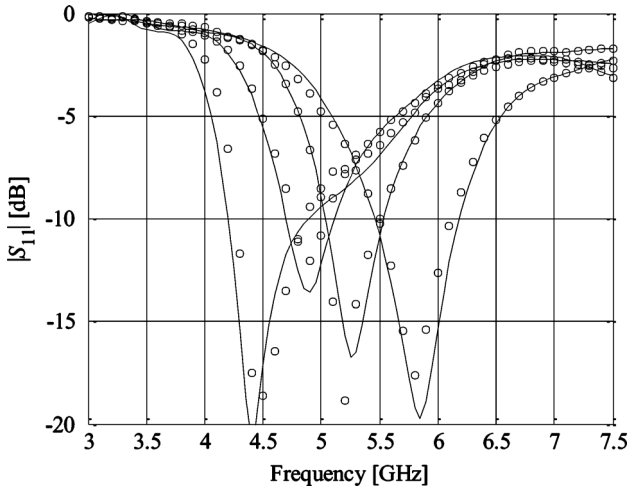


Fig. 7. Rectangular DRA: responses of $\boldsymbol{R}_f$ (—) and co-Kriging surrogate (o) at selected test points. Co-Kriging model created using 50 evaluations of $\boldsymbol{R}_f$ and 400 evaluations of $\boldsymbol{R}_c$.

TABLE II
RECTANGULAR DRA: MODELING RESULTS

Model	Average Modeling Error [%]				
	$N_{KR}=20$	$N_{KR}=50$	$N_{KR}=100$	$N_{KR}=200$	$N_{KR}=400$
$\boldsymbol{R}_{s.KR}$	12.1	8.8	6.9	5.2	3.6
$\boldsymbol{R}_{s.CO}$	6.7	5.4	5.0	4.1	3.5

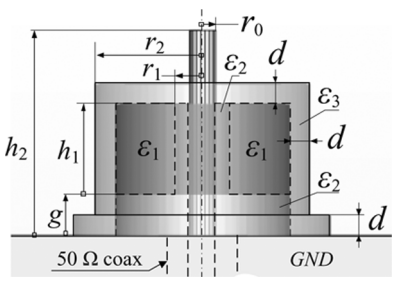


Fig. 8. Wideband hybrid antenna: geometry.

and $N_{\mathrm{KRf}} = 50$. The UWB dipole antenna was optimized starting from $\boldsymbol{x}^{\mathrm{init}} = [19\ 13\ 0.5\ 13\ 5\ 1]^T$ mm, using the following design specifications: $|S_{11}| \leq -12$ d B for 3.1 to 10.6 GHz. Fig. 10 shows the high-fidelity and co-Kriging model response at $\boldsymbol{x}^{\mathrm{init}}$, as well as the high-fidelity model response at the optimized co-Kriging model design $\boldsymbol{x}^* = [18.15\ 12.88\ 0.38\ 12.93\ 6.00\ 1.19]^T$ mm.

The rectangular DRA was also optimized using the co-Kriging model assuming the following design goal: $|S_{11}| \leq -10$ dB for 5 to 6 GHz. The optimization starts from $\boldsymbol{x}^{\mathrm{init}} = [10\ 8.5\ 0.5\ 3\ 10]^T$ mm. Fig. 11 shows the high-fidelity and co-Kriging model reflection response at $\boldsymbol{x}^{\mathrm{init}}$, as well as the high-fidelity model response at the optimized co-Kriging model design $\boldsymbol{x}^* = [9.53\ 8.10\ 0.90\ 2.64\ 11.31]^T$ mm. Fig. 12 shows the radiation response of the fine model at the final design.

Finally, the wideband hybrid DRA of Section III-C was optimized using the co-Kriging surrogate with respect to the following specifications: $|S_{11}| \leq -20$ dB for 8 to 13 GHz. The optimization starts from $\boldsymbol{x}^{\mathrm{init}} = [4\ 10\ 2\ 4\ 0.5]^T$ mm. The optimized design obtained using co-Kriging model is $\boldsymbol{x}^* = [3.99\ 9.93\ 1.63\ 3.55\ 0.19]^T$ mm (Fig. 13).

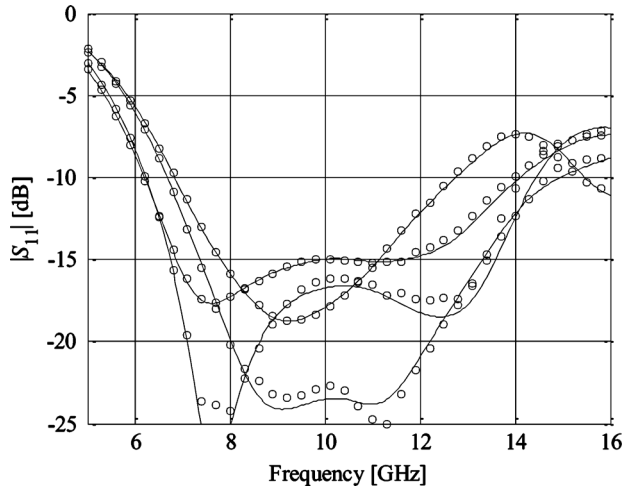


Fig. 9. Wideband hybrid DRA: responses of $\boldsymbol{R}_f$ (—) and co-Kriging surrogate (o) at selected test points. Co-Kriging model created using 50 evaluations of $\boldsymbol{R}_f$ and 400 evaluations of $\boldsymbol{R}_c$.

TABLE III
WIDEBAND HYBRID DRA: MODELING RESULTS

Model	Average Modeling Error [%]				
	$N_{KR}=20$	$N_{KR}=50$	$N_{KR}=100$	$N_{KR}=200$	$N_{KR}=400$
$\boldsymbol{R}_{s.KR}$	12.0	6.9	4.7	3.3	2.1
$\boldsymbol{R}_{s.CO}$	5.9	3.5	3.2	2.2	1.9

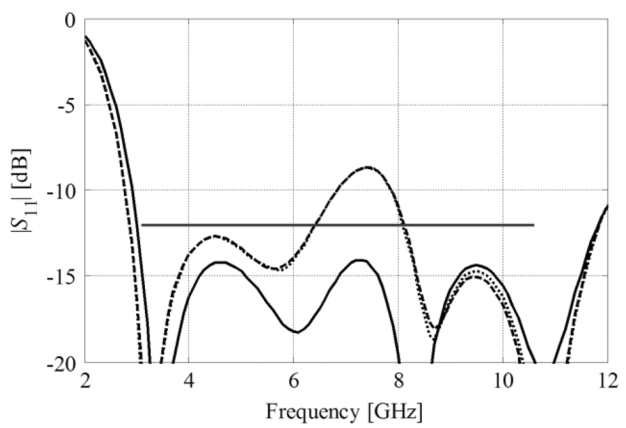


Fig. 10. Optimization of UWB dipole antenna using the co-Kriging model: responses of the high-fidelity (- - -) and the co-Kriging model ($\cdots$) at $\boldsymbol{x}^{\mathrm{init}}$, and the response of the high-fidelity model at the co-Kriging model optimum (—). The co-Kriging model has been created using 400 low-fidelity and 50 high-fidelity model samples (i.e., $N_{\mathrm{KRc}} = 400$ and $N_{\mathrm{KRf}} = 50$).

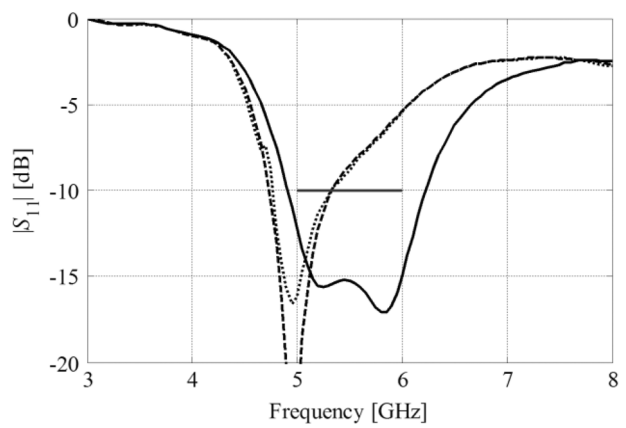


Fig. 11. Optimization of rectangular DRA using the co-Kriging model with $N_{\mathrm{KRc}} = 400$ and $N_{\mathrm{KRf}} = 50$: High-fidelity (- - -) and co-Kriging model ($\cdots$) responses at $\boldsymbol{x}^{\mathrm{init}}$, and high-fidelity model response at the co-Kriging model optimum (—).

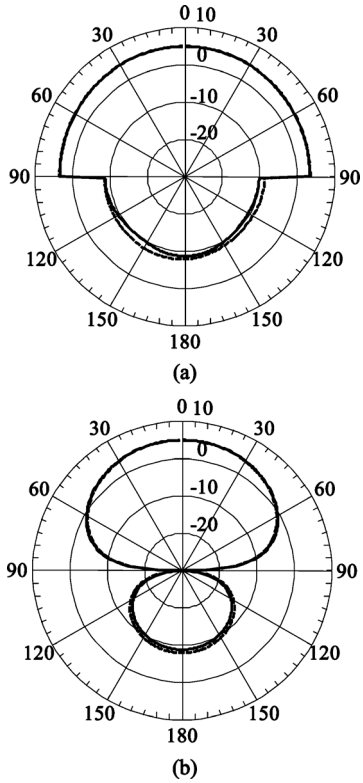


Fig. 12. Gain [dB] of the final DRA design at 5.00 (- -) and 6.00 (-) GHz: (a) co-pol. in the E-plane (marked by the horizontal dashed–dotted line in Fig. 5(b)), the feeding microstrip is on the left and (b) x-pol. in the H-plane [marked by the vertical dashed–dotted line in Fig. 5(b)]. Discontinuity of the gain pattern in (a) is due to the ground plane, which is modeled with infinite lateral dimensions.

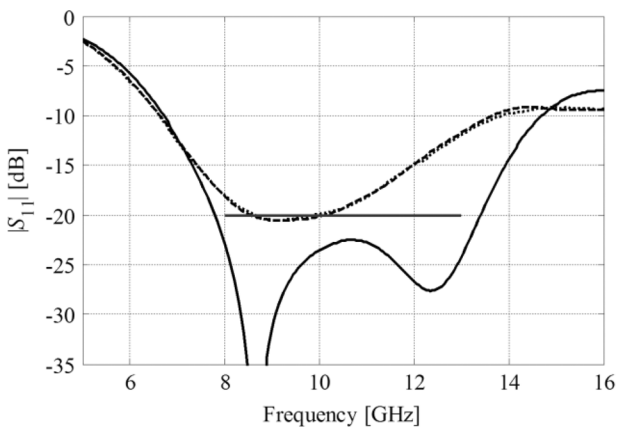


Fig. 13. Optimization of wideband hybrid DRA using the co-Kriging model with $N_{\mathrm{KRc}} = 400$ and $N_{\mathrm{KRf}} = 50$: High-fidelity (- - -) and co-Kriging model ($\cdots$) responses at $\boldsymbol{x}^{\mathrm{init}}$, and high-fidelity model response at the co-Kriging model optimum (—).

All three examples demonstrate that the co-Kriging model created with a relatively small number of high-fidelity model samples (only 50) can be reliably used to carry out parametric optimization of the antenna structures.

V. Conclusion

An antenna modeling methodology using co-Kriging has been presented. We have demonstrated that by combining the low- and high-fidelity EM simulations, it is possible to create an accurate antenna model using a limited number of high-fidelity data points. A typical number of data points a co-Kriging antenna model is configured with is substantially smaller than that needed for a conventional approximation technique relying exclusively on high-fidelity data sampling. Applications of co-Kriging models for antenna design are also discussed.

Acknowledgment

The authors would like to thank Computer Simulation Technology AG, Darmstadt, Germany, for making the CST Microwave Studio available.

References

[1] R. L. Haupt, "Antenna design with a mixed integer genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 577–582, Mar. 2007.
[2] A. J. Kerkhoff and H. Ling, "Design of a band-notched planar monopole antenna using genetic algorithm optimization," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 604–610, Mar. 2007.
[3] N. Jin and Y. Rahmat-Samii, "Parallel particle swarm optimization and finite- difference time-domain (PSO/FDTD) algorithm for multiband and wide-band patch antenna designs," *IEEE Trans. Antennas Propag.*, vol. 53, no. 11, pp. 3459–3468, Nov. 2005.
[4] A. Halehdar, D. V. Thiel, A. Lewis, and M. Randall, "Multiobjective optimization of small meander wire dipole antennas in a fixed area using ant colony system," *Int. J. RF Microw. CAE*, vol. 19, no. 5, pp. 592–597, 2009.
[5] A. I. J. Forrester and A. J. Keane, "Recent advances in surrogate-based optimization," *Prog. Aerosp. Sci.*, vol. 45, no. 1–3, pp. 50–79, Jan. –Apr. 2009.
[6] M. D. Buhmann and M. J. Ablowitz, *Radial Basis Functions: Theory and Implementations*. Cambridge, U.K.: Cambridge Univ., 2003.
[7] T. W. Simpson, J. Peplinski, P. N. Koch, and J. K. Allen, "Metamodels for computer-based engineering design: Survey and recommendations," *Eng. Comput.*, vol. 17, no. 2, pp. 129–150, July 2001.
[8] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidynathan, and P. K. Tucker, "Surrogate-based analysis and optimization," *Progr. Aerosp. Sci.*, vol. 41, no. 1, pp. 1–28, Jan. 2005.
[9] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, Aug. 2004.
[10] J. Meng and L. Xia, "Support-vector regression model for millimeter wave transition," *Int. J. Infrared Milimeter Waves*, vol. 28, no. 5, pp. 413–421, May 2007.
[11] L. Xia, J. Meng, R. Xu, B. Yan, and Y. Guo, "Modeling of 3-D vertical interconnect using support vector machine regression," *IEEE Microw. Wireless Compon. Lett.*, vol. 16, no. 12, pp. 639–641, Dec. 2006.
[12] M. Martinez-Ramon and C. Christodoulou, "Support vector machines for antenna array processing and electromagnetics," *Synthesis Lectures on Comput. Electromagn.*, vol. 1, no. 1, 2006.
[13] V. Miraftab and R. R. Mansour, "EM-based microwave circuit design using fuzzy logic techniques," in *Proc. Inst. Elect. Eng.—Microw., Antennas, Propag.*, Dec. 2006, vol. 153, no. 6, pp. 495–501.
[14] J. Zhai, J. Zhou, L. Zhang, and W. Hong, "Behavioral modeling of power amplifiers with dynamic fuzzy neural networks," *IEEE Microw. Wireless Compon. Lett.*, vol. 20, no. 9, pp. 528–530, 2010.
[15] J. E. Rayas-Sánchez, "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Trans. Microw. Theory Tech.*, vol. 52, no. 1, pp. 420–435, Jan. 2004.
[16] H. Kabir, Y. Wang, M. Yu, and Q. J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 4, pp. 867–879, Apr. 2008.
[17] Y. Cao, X. Chen, and G. Wang, "Dynamic behavioral modeling of nonlinear microwave devices using real-time recurrent neural network," *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1020–1026, 2009.
[18] H. Kabir, Y. Wang, M. Yu, and Q. J. Zhang, "High-dimensional neural-network technique and applications to microwave filter modeling," *IEEE Trans. Microwave Theory Tech.*, vol. 58, no. 1, pp. 145–156, Jan. 2010.
[19] G. S. A. Shaker, M. H. Bakr, N. Sangary, and S. Safavi-Naeini, "Accelerated antenna design methodology exploiting parameterized Cauchy models," in *Proc. PIER-99*, 2009, pp. 279–309.
[20] J. W. Bandler, N. Georgieva, M. A. Ismail, J. E. Rayas-Sánchez, and Q. J. Zhang, "A generalized space mapping tableau approach to device modeling," *IEEE Trans. Microw. Theory Tech.*, vol. 49, no. 1, pp. 67–79, Jan. 2001.
[21] E. Rayas-Sánchez and V. Gutierrez-Ayala, "EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 12, pp. 4528–4537, 2006.

[22] S. Koziel and J. W. Bandler, "Recent advances in space-mapping-based modeling of microwave devices," *Int. J. Numer. Modell.*, vol. 23, no. 6, pp. 425–446, 2010.

[23] L. Zhang, Q. J. Zhang, and J. Wood, "Statistical neuro-space mapping technique for large-signal modeling of nonlinear devices," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 1, pp. 2453–2467, 2011.

[24] J. W. Bandler, Q. S. Cheng, and S. Koziel, "Simplified space mapping approach to enhancement of microwave device models," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 16, no. 5, pp. 518–535, 2006.

[25] S. Koziel, J. W. Bandler, and K. Madsen, "Theoretical justification of space-mapping-based modeling utilizing a data base and on-demand parameter extraction," *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 12, pp. 4316–4322, Dec. 2006.

[26] S. Koziel and J. W. Bandler, "A space-mapping approach to microwave device modeling exploiting fuzzy systems," *IEEE Trans. Microw. Theory Tech.*, vol. 55, no. 12, pp. 2539–2547, Dec. 2007.

[27] S. Koziel and J. W. Bandler, "Accurate modeling of microwave devices using space mapping and Kriging," in *Int. Rev. Progr. Appl. Comput. Electromagn.*, Tampere, Finland, Apr. 26–29, 2010, pp. 902–907.

[28] M. C. Kennedy and A. O'Hagan, "Predicting the output from complex computer code when fast approximations are available," *Biometrika*, vol. 87, pp. 1–13, 2000.

[29] S. Ogurtsov and S. Koziel, "Simulation-driven design of dielectric resonator antenna with reduced board noise emission," in *IEEE MTT-S Int. Microw. Symp. Dig.*, 2011, pp. 1–4.

[30] J. Sacks, W. J. Welch, T. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statist. Sci.*, vol. 4, no. 4, pp. 409–435, 1989.

[31] D. Gorissen, K. Crombecq, I. Couckuyt, P. Demeester, and T. Dhaene, "A surrogate modeling and adaptive sampling toolbox for computer based design," *J. Mach. Learn. Res.*, vol. 11, pp. 2051–2055, 2010.

[32] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

[33] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.*, vol. 13, pp. 455–492, 1998.

[34] I. Couckuyt, F. Declercq, T. Dhaene, and H. Rogier, "Surrogate-based infill optimization applied to electromagnetic problems," *Adv. Design Optim. Microw./RF Circuits Syst.*, vol. 20, p. 492, 2010.

[35] G. Wang and S. Shan, "Review of metamodeling techniques in support of engineering design optimization," *J. Mechan. Des., ASME*, vol. 129, pp. 370–380, 2007.

[36] J. Staum, "Better simulation metamodeling: The why, what, and how of stochastic kriging," in *Proc. 2009 Winter Simulation Conf.*, 2009, pp. 119–133.

[37] M. D. Morris, T. J. Mitchell, and D. Ylvisaker, "Design and analysis of computer experiments: Use of derivatives in surface prediction," *Technometrics*, vol. 35, no. 3, pp. 243–255, 1993.

[38] A. I. Forrester, A. Sobester, and A. J. Keane, "Multi-fidelity optimization via surrogate modelling," in *Proc. Roy. Soc.*, 2007, vol. 463, pp. 3251–3269.

[39] B. Beachkofski and R. Grandhi, "Improved distributed hypercube sampling," Amer. Inst. of Aeronautics Astronautics, paper AIAA 2002-1274, 2002.

[40] T. G. Spence and D. H. Werner, "A novel miniature broadband/multiband antenna based on an end-loaded planar open-sleeve dipole," *IEEE Trans. Antennas Propag.*, vol. 54, no. 12, pp. 3614–3620, Dec. 2006.

[41] CST Microwave Studio. ver. 2011, CST AG, Darmstadt, Germany, 2011.

[42] A. Petosa, *Dielectric Resonator Antenna Handbook*. Norwood, MA, USA: Artech House, 2007.

[43] J. Kleijnen, *Design and Analysis of Simulation Experiments*. New York, NY, USA: Springer, 2008.

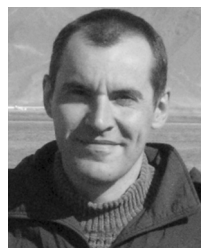
[44] D. Krige, "A statistical approach to some basic mine valuation problems on the Witwatersrand," *J. Chem., Metallurgic., Mining Soc. South Africa*, vol. 52, pp. 119–139, 1951.

[45] M. T. M. Emmerich, K. Giannakoglou, and B. Naujoks, "Single- and multiobjective evolutionary optimization assisted by Gaussian random field metamodels," *IEEE Trans. Evolution. Comput.*, vol. 10, pp. 421–439, 2006.

[46] I. Couckuyt, A. Forrester, D. Gorissen, F. De Turck, and T. Dhaene, "Blind Kriging: Implementation and performance analysis," *Adv. Eng. Softw.*, vol. 49, pp. 1–13, 2012.

Slawomir Koziel (M'03–SM'07) received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1995 and 2000, respectively, and the M.Sc. degrees in theoretical physics and in mathematics, in 2000 and 2002, respectively, as well as the Ph.D. in mathematics, in 2003 from the University of Gdansk, Gdansk, Poland.

He is currently a Professor with the School of Science and Engineering, Reykjavik University, Iceland. His research interests include CAD and modeling of microwave circuits, simulation-driven design, surrogate-based optimization, space mapping, circuit theory, analog signal processing, evolutionary computation, and numerical analysis.



Stanislav Ogurtsov received the Physicist degree from Novosibirsk State University, Novosibirsk, Russia, in 1993, and the Ph.D. degree in electrical engineering from Arizona State University, Tempe, AZ, USA, in 2007.

He is currently a Postdoctoral Researcher at the Electromagnetic Optimization and Modeling Center, Reykjavik University, Iceland. His research interests include simulation-driven computer aided design of RF, microwave, and millimeter-wave circuits, ultra-wideband antennas, computational electromagnetics, modeling of high-speed digital circuits, and material characterization.

Ivo Couckuyt received the M.Sc. degree in computer science from the University of Antwerp (UA), Antwerp, Belgium, in 2007.

In October 2007, he joined the research group Computer Modeling and Simulation (COMS) (now merged with CoMP), supported by a research project of the Fund for Scientific Research Flanders (FWO-Vlaanderen). Since January 2009, he has been working towards the Ph.D. degree in the research group INTEC Broadband Communication Networks (IBCN) at Ghent University, Ghent, Belgium. At the same time, he remains affiliated with CoMP. His research activities include surrogate modeling, surrogate-based optimization, and inverse problems of time-consuming problems.

Tom Dhaene (SM'06) received the M.Sc. degree and the Ph.D. degree in electrical engineering from Ghent University, Ghent, Belgium, in 1989 and 1993, respectively.

Since October 2000, he has been a Professor in the Computer Modeling and Simulation (COMS) research group in the Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium. Since October 2007, he has been a Professor in the Faculty of Engineering in the INTEC Broadband Communication Networks (IBCN) research group of the Department of Information Technology (INTEC), Ghent University, Ghent, Belgium. His research interests include distributed scientific computing, machine learning, bioinformatics, signal integrity, electromagnetic compatibility, model order reduction, optimal design, surrogate modeling (or metamodeling) of complex systems, circuit and EM modeling of high-speed interconnections and broadband communication systems, adaptive system identification of deterministic LTI systems, and numerical analysis techniques.